DOCUMENT RESUME

ED 054 657 FL 002 354

AUTHOR Oh, Choon-Kyu

TITLE Spurious Counterexamples to the Complex NP Constraint, Including a "Variably Crazy Rule" in Korean Syntax.

INSTITUTION Hawaii Univ., Honolulu. Dept. of Linguistics.

PUB DATE Dec 70

NOTE 20p.; In Working Papers in Linguistics, v2 n9 p105-124 Dec 1970

EDRS PRICE MF-$0.65 HC-$3.29

DESCRIPTORS Deep Structure; Descriptive Linguistics; English; *Korean; Language Universals; *Nominals; *Phrase Structure; Sentences; Sentence Structure; *Structural Analysis; Structural Grammar; Syntax; Transformation Generative Grammar; Transformations (Language); *Transformation Theory (Language)

ABSTRACT

In order to make linguistically significant generalizations in formulating transformational rules, a limited number of variables must be used. Such a limitation rule is the Complex Noun Phrase (NP) Constraint, active in English and, according to claims, universal. The fact that Korean synta, which is greatly different from English, also requires such a constraint supports the claim of the universality of this rule. Apparent violations of the Complex NP Constraint in Korean may be attributed to two successive processes, which, when combined, do not contradict the rule. Rule-simplification, one of the major forces in linguistic change, is at work in the application of rules here. A list of references is provided. (VM)

ED054657

From: Working Papers in Linguistics; Vol. 2,
No. 9, December 1970.

SPURIOUS COUNTEREXAMPLES TO THE COMPLEX NP CONSTRAINT, INCLUDING A "VARIABLY CRAZY RULE" IN KOREAN SYNTAX[1]

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

Choon-Kyu Oh

University of Hawaii

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY Choon-Kyu Oh

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

In formulating transformational rules, the utilization of variables is inescapable. But the use of variables must be limited. Otherwise, many linguistically significant generalizations cannot be captured, and also the derivation of some ill-formed sentences cannot be blocked. As one of the necessary constraints, Ross (1967: sec. 4.20) set up what he called the Complex NP Constraint (hence-after CNC), which he defined thus:

1) The Complex NP Constraint: No element contained in a sentence dominated by a noun phrase with a lexical head noun may be moved out of that noun phrase by a transformation.

This constraint is needed to block such sentences as (2), but allow those like (3).

2a) *Who does Phineas know a girl (who is) working with?

b) *The hat which I believed the claim that Otto was wearing is red.

3a) Who does Phineas know that is working with the girl?

b) The hat which I believed that Otto was wearing is red.

Ross (1967: sec. 4.5) claimed, further, that (1) is universal. The purpose of this paper is to show the plausibility of this claim by revealing that Korean--which is so far apart from English typologically, areally, and genetically--also needs it.

FL002354

RELATIVIZATION in Korean will be tentatively formulated as (4), disregarding all aspects which are irrelevant to this discussion:

4) RELATIVIZATION:

$$\underbrace{W, \quad [X,}_{1} \quad NP_i, \quad 2 \quad Y]_S, \quad 3 \quad NP_i, \quad 4 \quad Z \quad 5 \Longrightarrow 1, \emptyset, 3+rel, 4, 5$$

where rel is the relativizer -n.

The (a) sentences of (5) would be relativized to (b) sentences, respectively, by (4):

5.1a) [Mary-ka ilyoil-ey Seoul-eyse kù si-lúl ssúta]$_S$ Mary
Sunday the poem-obj wrote
'[Mary wrote the poem in Seoul on Sunday]$_S$ Mary'

b) ____ ilyoil-ey Seoul-eyse kù si-lúl ssú-n Mary
'Mary who wrote the poem in Seoul on Sunday'

.2a) [Mary-ka ilyoil-ey Seoul-eyse kù si-lúl ssúta]$_S$ ilyoil

b) Mary-ka _____ Seoul-eyse kù si-lúl ssú-n ilyoil
'The Sunday on which Mary wrote the poem in Seoul'

.3a) [Mary-ka ilyoil-ey Seoul-eyse kù si-lúl ssúta]$_S$ Seoul

b) Mary-ka ilyoil-ey _____ kù si-lúl ssú-n Seoul
'Seoul where Mary wrote the poem on Sunday'

.4a) [Mary-ka ilyoil-ey Seoul-eyse kù si-lúl ssúta]$_S$ si

b) Mary-ka ilyoil-ey Seoul-eyse _____ ssú-n si
'The poem which Mary wrote in Seoul on Sunday'

Before we proceed to apparent counterexamples to the CNC, PRONOMINALIZATION, which is not subject to the CNC or any other of Ross's constraints, should be distinguished from RELATIVIZATION. The distinction will be made clear below. In Korean, PRONOMINALIZATION deletes all the correferential NP's except the first one

in the matrix sentence. Langacker's (1969) constraint on PRONOMINALIZATION seems to hold here. That is, the pronominalized NP cannot both command and precede the antecedent. Also all the notions utilized in the output condition on PRONOMINALIZATION by Lakoff (1968) appear to be relevant here, too. Under conditions which are irrelevant to our arguments and are not given here, this rule, PRONOMINALIZATION becomes obligatory. Thus, the (a) sentences of (6) are pronominalized to the respective (b) sentences:

6.1a) John$_i$-nin [John$_i$-i Shakespeare ilk]$_S$-ki-lil cohahanta
read nominal-obj likes

b) John-nin _____ Shakespeare ilk-ki-lil cohahanta

'John likes to read Shakespeare.'

.2a) Mary$_i$-nin [Mary$_i$-ka John-lil cuki-n]$_S$ hyemiy-lil patkoissta
killed suspicion receives

b) Mary-nin _____ John-lil cuki-n hyemiy-lil patkoissta

'Mary is suspected to have killed John.'

.3a) [John$_i$-i ki yenghwa-lil po-n]$_S$ cek-i John$_i$-eyke-nin issta
the movie watched-experience is

b) _____ ki yenghwa-lil po-n cek-i John-eyke-nin issta

'John has watched the movie.'

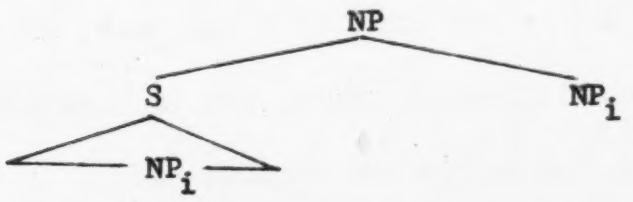
.4a) John$_i$-nin [John$_i$-i sayngkakhay o-ten kes]$_S$-lil cekessta
have been thinking thing wrote

b) John-nin $\left(\begin{Bmatrix} \text{ki} \\ \text{caki} \end{Bmatrix}\text{-ka}\right)$ sayngkakhay o-ten kes-lil cekessta

'John wrote down what he had been thinking of.'

That the deletion of NP's in (6) is PRONOMINALIZATION rather than something equivalent to the English EQUI-NP DELETION is clearly demonstrated in the optional nature of the pronouns in (6.4b).

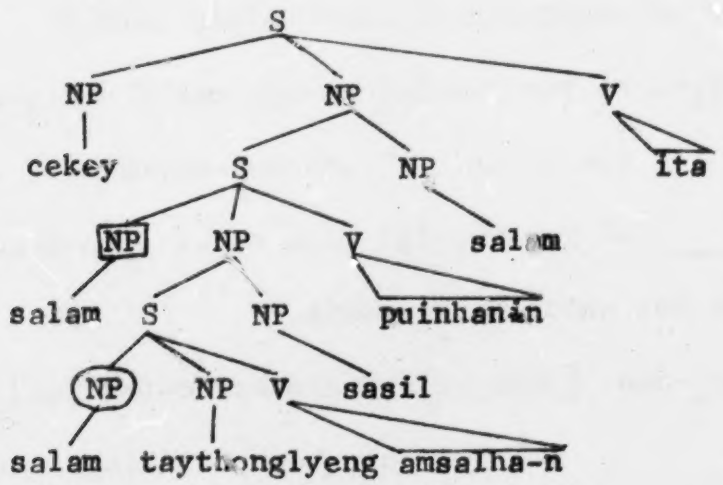
The structural difference between PRONOMINALIZATION and RELATIVIZATION is rather obvious. Only in the latter is it true that the deletion of an embedded NP is conditioned by its correferentiality with the head noun which is schematically shown below, while in the former, the antecedent is not the head noun.



Therefore, a sentence like (7) should not be considered to be a counterexample to the CNC. In what follows, each (b) tree represents the underlying structure of the respective (a) sentence.

7.1a) cekey taythonglyeng-lǎl amsalha-n sasil-lǎl puinhanǎn salamita
that President assassinated fact denying man is
'That is the man who denies the fact that (he) assassinated the President.'

b)



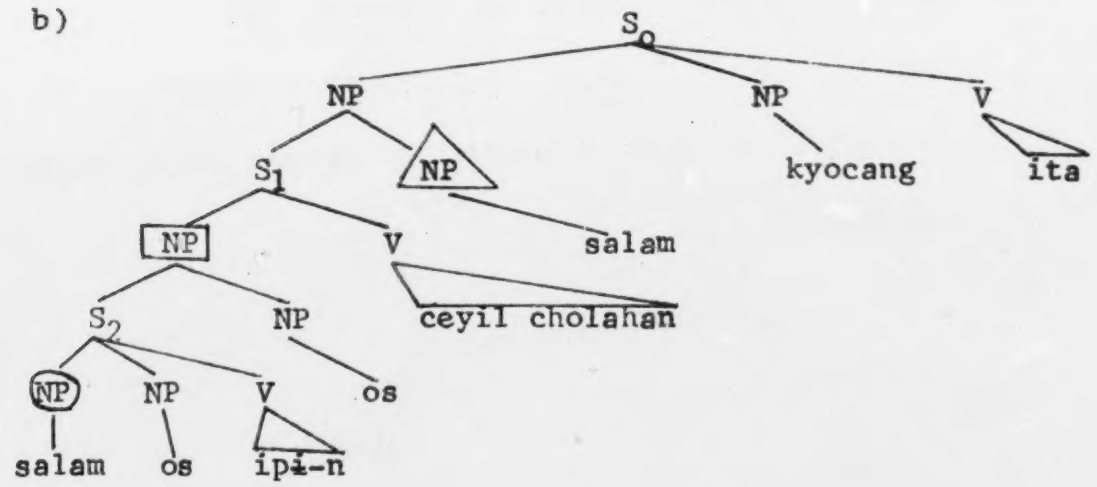
.2) cekey caki-ka taythonglyeng-lǎl amsalha-n sasil-lǎl puin
he
hanǎn salamita.
'That is the man who denies the fact that he assassinated the President.'

In (7), the circled NP is deleted by PRONOMINALIZATION under correferentiality with the boxed NP, which in turn is deleted by RELATIVIZATION. That the deletion of the circled NP here is by PRONOMINALIZATION is again demonstrated by the possibility of (7.2) and the non-correferential head noun.

Now, consider the following:

8a) ipi-n os-i ceyil cholaha-n salam-i kyocang ita
wear clothes most shabby man principal is
'The man with the shabbiest clothes on is the principal.'

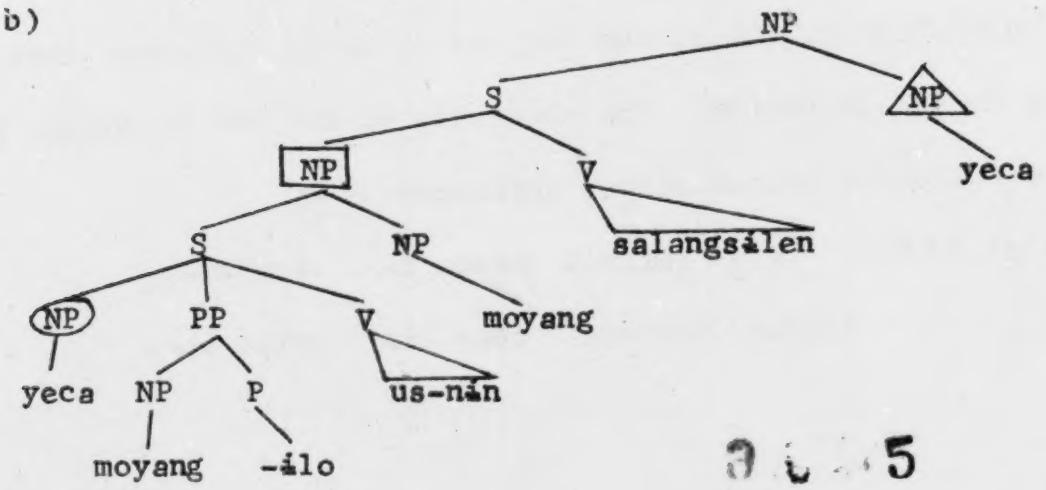
b)



Here, os 'clothes' in S_2 is deleted by RELATIVIZATION and the deletion of salam 'man' there will be explained later.

9a) us-nin moyang-i salangsile-n yeca
smile way attractive girl
'girl who smiles attractively'

b)

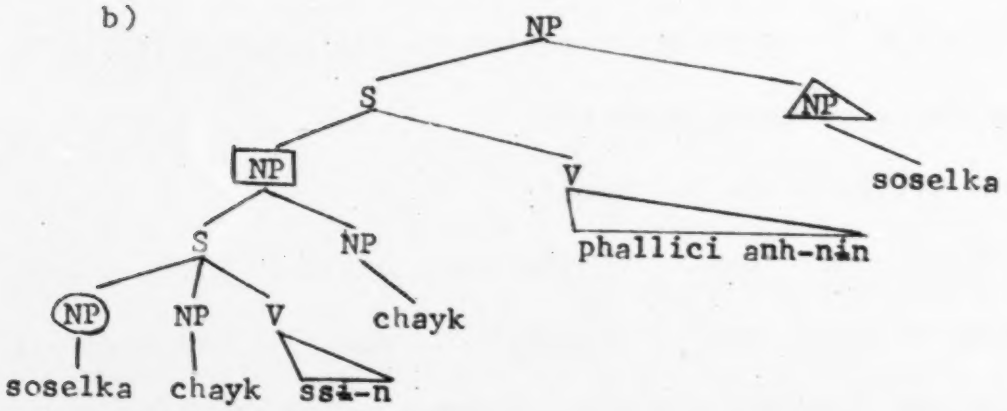


10a) ssi-n chayk-i phallici anh-nin soselka

wrote book doesn't sell well novelist

'the novelist whose book does not sell well'

b)

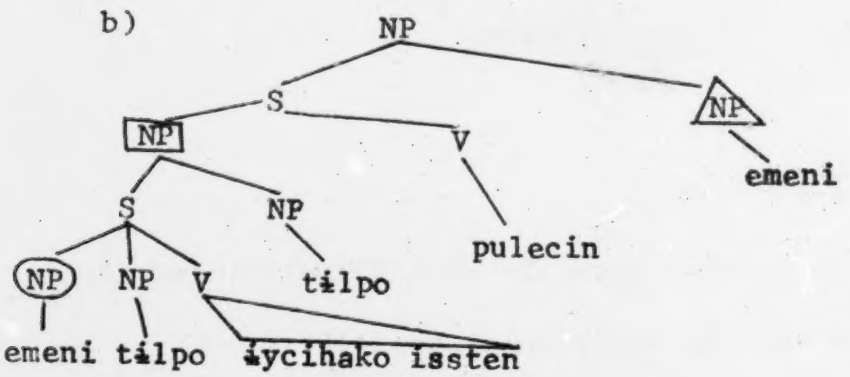


11a) iycihako isste-n tilpo-ka puleci-n emeni

lean on beam is broken mother

Lit. 'Mother whose beam is broken' = 'Mother whose mainstay is gone.'

b)



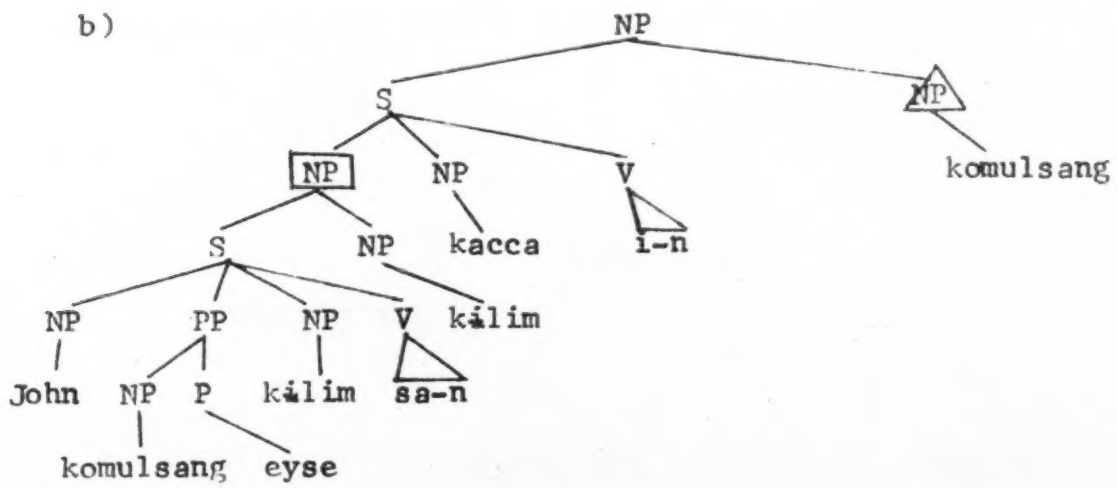
The deletion of the circled NP's appears to be a violation of the CNC, since they were within the scope of the boxed complex NP's and no antecedents precede them to justify regarding the deletion as PRONOMINALIZATION. This, however, by no means indicates that the CNC is not needed in Korean. The necessity of the CNC in Korean is shown by the following ungrammatical sentences:

12a) *John-i sa-n kilim-i kacca i-n komulsang

bought picture fake is curio shop

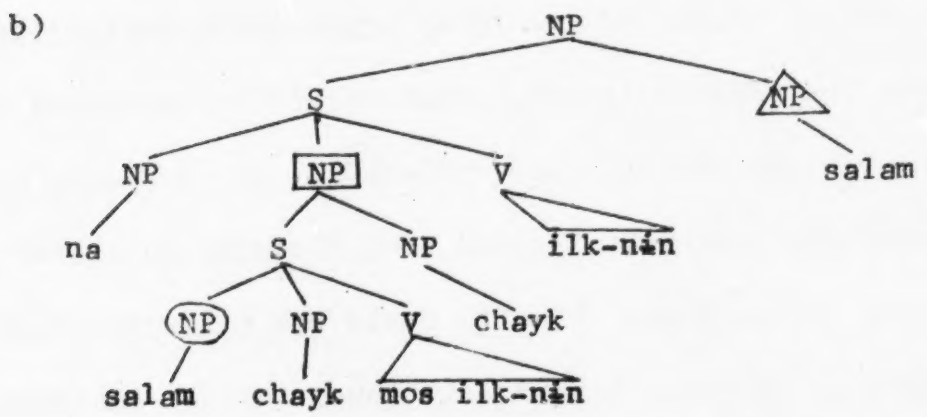
12a) *'the curio shop where the picture which John bought (there)
was a counterfeit.'

b)



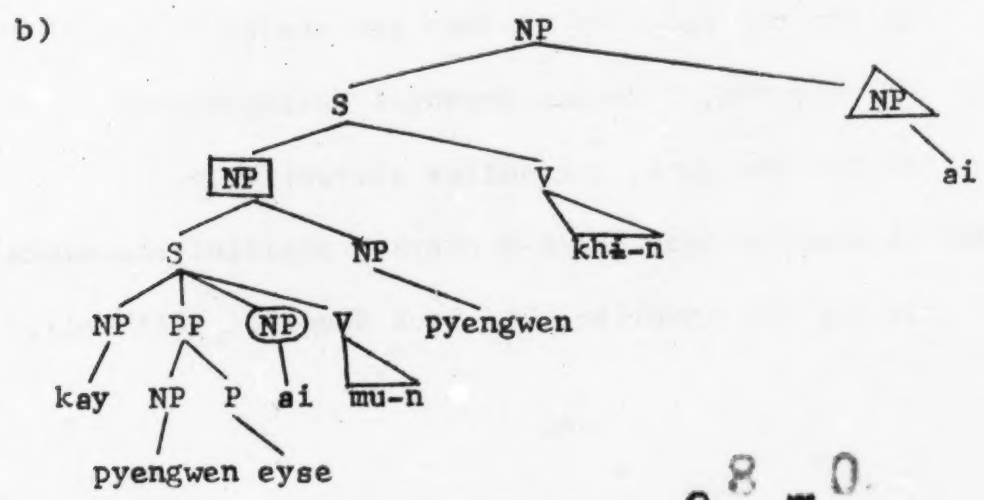
13a) *mos ilk-nin chayk-lil nay-ka ilk-nin salam
cannot read book I read man
*'the man who I can read the book which (he) cannot read'

b)



14a) *kay-ka mu-n pyengwen-i khi-n ai
dog bit hospital large child
*'the child who the hospital where a dog bit (him) is large'

b)



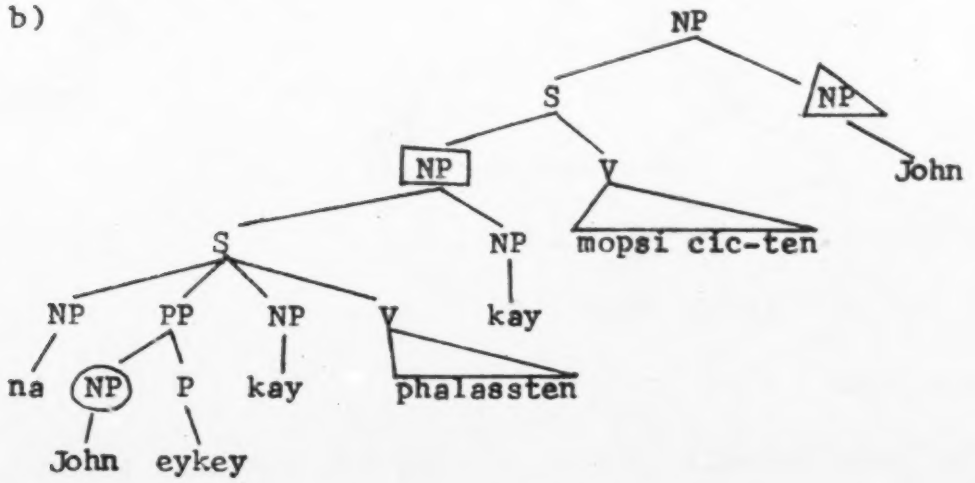
0 8 7 0

15a) *nay-ka phalass-ten kay-ka mopsi cic-ten John

I sold dog fiercely barked

*'John who the dog which I sold (him) barked fiercely'

b)



The ungrammaticality of the preceding is due to the violation of the CNC occasioned by deleting the circled NP's. These NP's were within the scope of the boxed complex NP's; again PRONOMINALIZATION is not responsible for their deletion. What has to be explained is the difference between (8)-(11) and (12)-(15). All the structural differences have been investigated, but the only conceivable difference is that only in sentences (8)-(11) could the triangled head noun be the topic of embedded sentences. Thus, the (c) sentences of (8)-(11) are grammatical while those of (12)-(15) are impossible.

8c) kǎ salam-nǎn, ipǎ-n os-i chalshata

topic

'As for the man, his clothes are shabby.'

9c) kǎ yeca-nǎn, us-nǎn moyang-i salangsǎlepta

'As for the girl, she smiles attractively.'

10c) kǎ soselka-nǎn, ssǎ-n chayk-i phallici anh-nǎnta

'As for the novelist, his book does not sell well.'

80 8

11c) emeni-nàn, àycihako issten tàlpo-ka pulecessta

'As for Mother, her mainstay is gone.'

12c) *kà komulsang-nàn, John-i sa-n kàlim-i kacca ita

*'As for the curio shop, the picture John bought (there) is a fake.'

13c) *kà salam-nàn, mos ilk-nàn chayk-làl nay-ka ilk-nànta

*'As for the man, I can read the book which (he) cannot read.'

14c) *kà ai-nàn, kay-ka mu-n pyengwen-i khàta

*'As for the child, the hospital where the dog bit (him) is large.'

15c) *John-nàn, nay-ka phallass-ten kay-ka mopsi cicessta

*'As for John, the dog which I sold (him) barked fiercely.'

Because only the relativized NP's of (8)-(11) are outside the boxed complex NP's when RELATIVIZATION applies, the CNC is not violated in deleting them, and so they are grammatical.[2] Kuno (MS) independently discovered a set of syntactic similarities between TOPICALIZATION and RELATIVIZATION in Japanese and claimed that only the topicalized NP's are relativized.

If, however, the CNC operates in Korean, why is TOPICALIZATION not subject to it? What makes TOPICALIZATION possible for the relativized NP's in (8)-(11), but blocks it in (12)-(15)? Unless these matters are accounted for, we have not really shown that the CNC is needed in Korean, since (8)-(11) would stand as the counterexamples only of a different nature, i.e. not to RELATIVIZATION, but this time to TOPICALIZATION.

That TOPICALIZATION is subject to the CNC in Korean is shown by the ungrammaticality of the (c) sentences of (12)-(15). What

then makes the (c) sentences of (8)-(11) grammatical? The POSS-INSERTION rule, which is formulated as (16) below, appears to be extremely wide-spread among languages of the world. This rule raises the embedded subject and inserts a possessive marker after it. See (17.2) and (18.2), whose underlying structures are given as (17.1) and (18.1), respectively:

16) POSS-INSERTION: $X, \underset{1}{\left[\left[\left[NP, \underset{2}{Y}\right]_S\right]_{NP}, \underset{3}{Z}\right]}_{\sim NP} \underset{4}{}$

$$\Longrightarrow 1, \left[\left[2 + poss\right]_{NP}, \left[\left[3\right]_S\right]_{NP}\right]_{NP}, 4$$

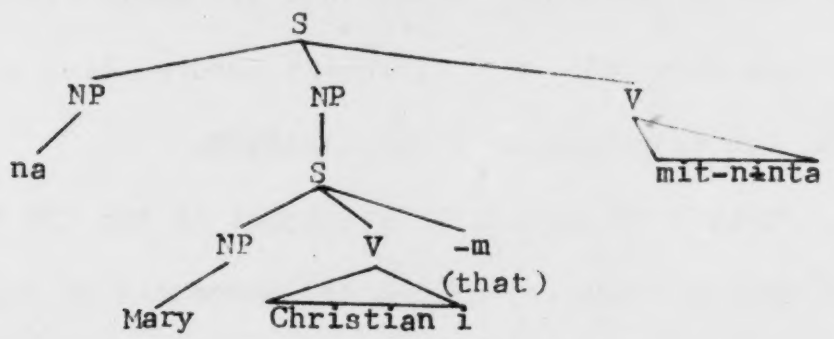
This rule is designed to show that at one stage of the history of Korean, POSS-INSERTION was permissible only for embedded sentences dominated by an NP which in turn is not directly dominated by another NP, i.e. complementation. This is still the case with languages like English. In formulating the rule, I assumed that all the embedded sentences are dominated by NP node, which point will be reflected by the phrase structures hereafter. (For the evidences of this, Sandra: 1970b p. 35). In this respect, RELATIVIZATION was formulated in traditional manner to avoid the unnecessary complication and all the trees were accordingly drawn up to this point.

17.1a) na-nin Mary-ka Christian i-m-lil mit-ninta

is-nom-obj believe

'I believe that Mary is a Christian.'

b)

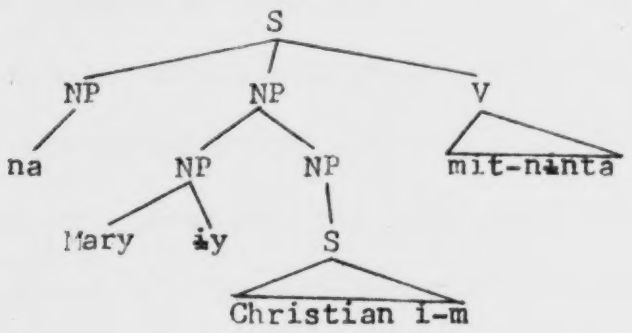


17.2a) na-nin Mary-iy Christian-i-m-lil mit-ninta

of

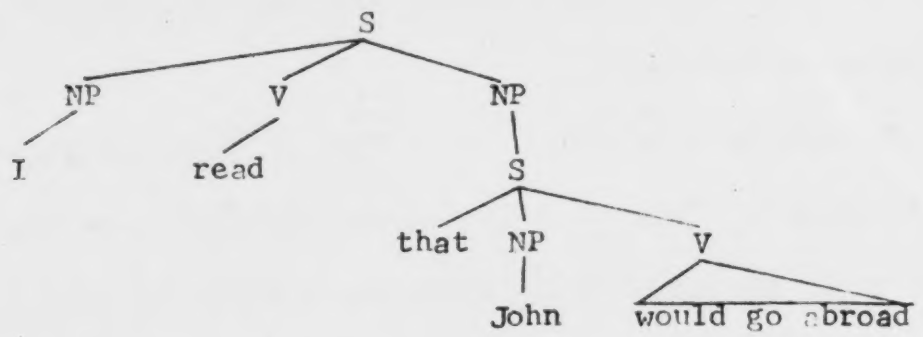
Lit. 'I believe Mary's being a Christian.'

b)



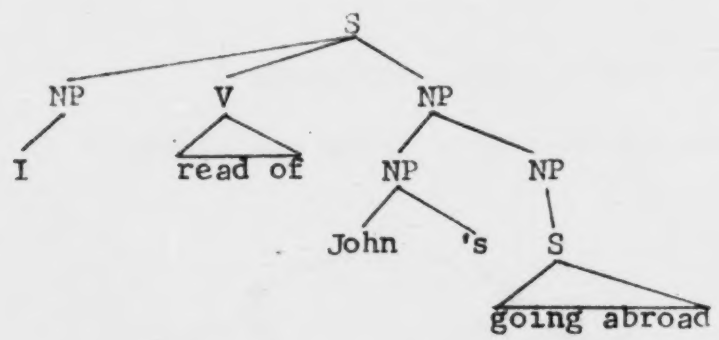
18.1a) I read that John would go abroad.

b)



.2a) I read of John's going abroad.

b)



Korean syntax, like Japanese, has been simplified by deleting the specification '$]_{\sim NP}$' of (16). In other words, it generalizes to almost all the embedded sentences, including relative clauses.[3] However, what has to be mentioned here is that when the head noun and the embedded sentence are in apposition, the rule applies much less, if ever, than when they are in the relative clause-antecedent

SI 110

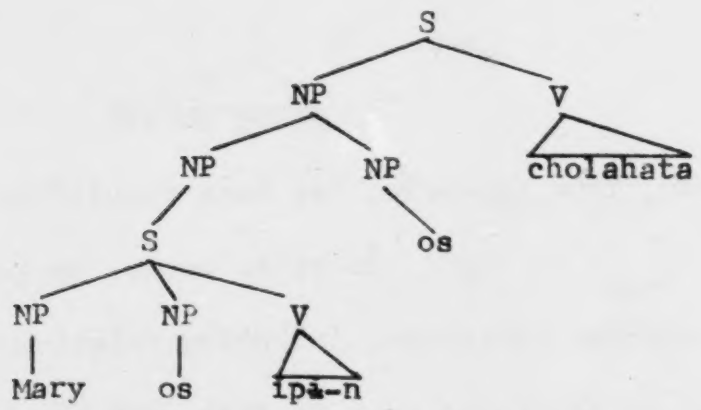
relation. This appears to show that the generalization of (16) is still in the variable stage and the variability depends on the feature matrix of the head noun included in Y, which is term A. Right now, this rule applies mostly to subject NP of an embedded sentence, and only marginally applicable to some objects, which seems to indicate the variable expansion of the rule to object NP's. Until sufficient statistics are available, the exact rule formulation with the variability incorporated will be postponed. (For variable rules, Labov: [1966], [1969], [1970]). For those speakers who allow Y to include appositive head nouns, (16) is applicable to the subject of any embedded sentence.

As will be made clear, this kind of generalization is not unnatural at all. This rule has become applicable to the embedded sentence under the complex NP (changing [19.1] to [19.2]) as well as normal nominalized sentence (changing [17.1] to [17.2], and [18.1] to [18.2]).

19.1a) Mary-ka ipa-n os-i cholahata

'The clothes that Mary wears are shabby.'

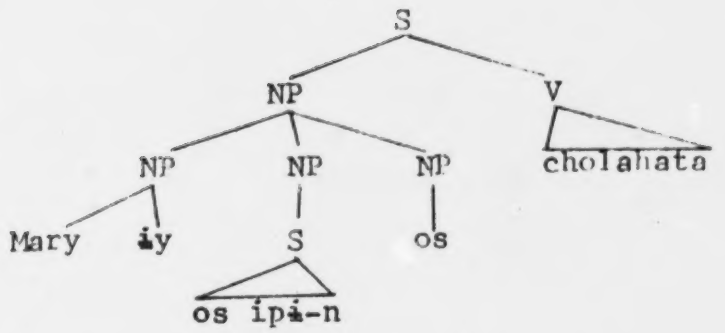
b)



.2a) Mary-ay ipa-n os-i cholahata

*'The clothes that Mary's wears are shabby.'

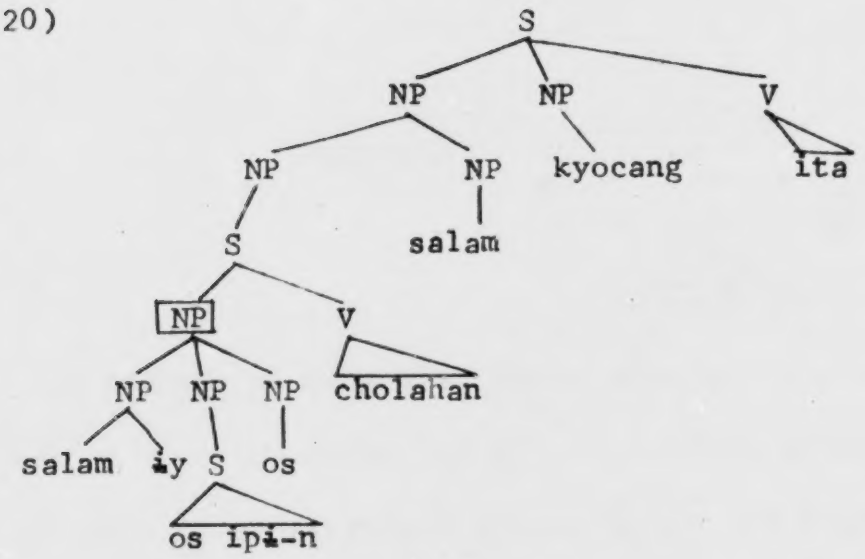
19.2b)



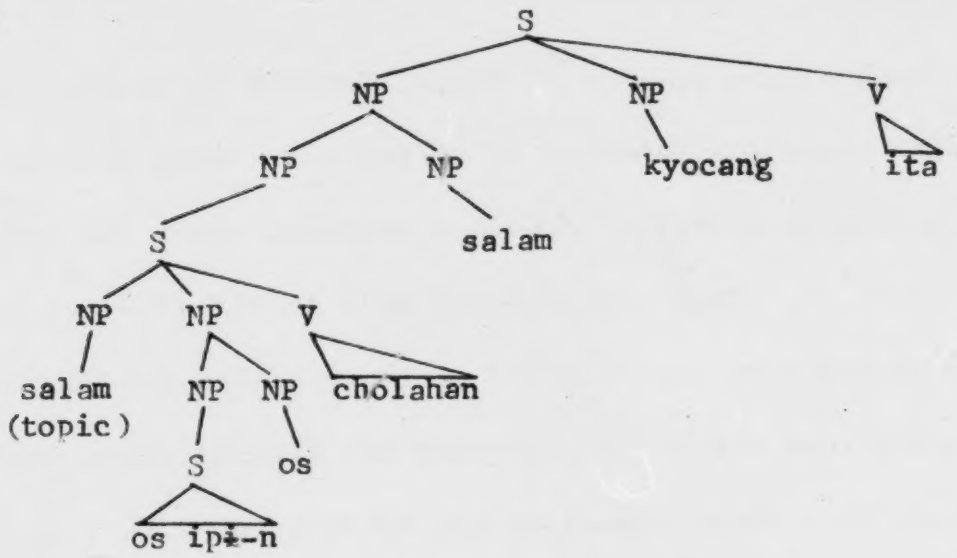
In (19), the CNC is not violated, since, although an element is moved from the embedded sentence, it is not taken out of the complex NP. See notes 2 and 3 for the necessary restrictions on rule (16) and TOPICALIZATION.

The generalized version of POSS-INSERTION is responsible for the apparently anomalous behavior of Korean with respective to the CNC. This serves as a bridge between a sentence under the complex NP and one above it. Thus, the deleted NP's in (8)-(11) are first raised or promoted from the embedded sentence within the complex NP and then topicalized out of that complex NP; finally these are deleted by RELATIVIZATION. These processes are illustrated with (8). After POSS-INSERTION, (8b) becomes (20), which in turn is changed into (21) by TOPICALIZATION. And then RELATIVIZATION gives (8a) from (21) by deleting the NP's salam and os under correferentiality with the antecedent--not directly from (8b), in which case the CNC would have been violated.

20)



21)



Notice that Ross's definition of the CNC leaves a loophole just in case there are two successive processes, the first of which moves an element out of the embedded sentence within the complex NP and the second of which moves that element out of the complex NP. This is precisely what has been done by POSS-INSERTION and TOPICALIZATION. POSS-INSERTION does not violate the CNC because it moves an element <u>within</u> the complex NP and not <u>out</u> of it. TOPICALIZATION does not violate the CNC because, although it moves an element which is in the complex NP, the NP to be topicalized is now out of the embedded sentence.

Our analysis has shown that neither RELATIVIZATION nor TOPICALIZATION has to violate the CNC in Korean, because of the generalized version of the POSS-INSERTION rule. The generalized variant of POSS-INSERTION may appear to be very unnatural to speakers of languages which have only the non-generalized form of the rule. Thus, (23), which is the English counterpart to Japanese (22b) and Korean (19.2), will sound odd to English speakers:

22a) Mary-ga kiteiru kimono
is wearing
'the kimono that Mary is wearing'

b) Mary-no kiteiru kimono
Poss

23a) *The clothes that Mary's is wearing

b) *Mary's wearing clothes (from 'the clothes that Mary is wearing')

Bach and Harms (MS) have shown for phonology that very implausible rules are created by rule-simplifications, illustrating their points with Japanese phonology. The generalized version of POSS-INSERTION in the present paper could be considered a syntactic parallel, even though, as Zwicky (1968) pointed out, naturalness is not as well defined in syntax as in phonology. Rule-simplification --which is one of the major forces in linguistic change--results in crazy rules in syntax as well as in phonology. Thus, however implausible it may sound to speakers of languages which allow the application of POSS-INSERTION only to nominalized sentences, the expansion of the rule application to any embedded sentence dominated by NP-node is a natural effect of rule-simplification.

NOTES

[1]I would like to express my deep appreciation to Professor C.-J. Bailey for his encouragement and stimulus in theoretical aspects and to Professor L. Josephs for insights on Japanese syntax, which parallels that of Korean in many significant ways.

[2]When the subject NP of the matrix verb is different from that of the embedded sentence, the applicability of TOPICALIZATION to the NP-Poss construction appears to depend on the type of the verb in the matrix sentence, as is shown below. (See my dissertation, for the specification of these verbs and the particle change nàn to ka in [2] below.)

1a) na-nàn kǎ yeca-ǎy us-nàn moyang-lǎl cohahanta
I the girl's smiling way like
'I like the girl's way of smiling.'

b) ?kǎ yeca-nàn, nay-ka us-nàn moyang-lǎl cohahanta
'As for the girl, I like her way of smiling.'

2a) na-nàn kǎ-ǎy mos ilk-nàn chayk-lǎl ilk-nánta
I his cannot read book read
*'I can read the book that his cannot read.'

b) *kǎ-nàn, nay-ka (kǎ-ka) mos ilknàn chayk-lǎl ilknánta
*'As for him, I can read the book which (he) cannot read.'

cohahanta 'like' is one example of a class of verbs including silhehata 'dislike', kiekhata 'remember' etc., which allows topicalization of the NP-Poss construction.

[3]To most of the speakers, the application of POSS-INSERTION to the embedded sentence appositive to its head noun is illegitimate. (For details, see my forthcoming dissertation [Oh, MS]). The (b) sentences which are derived from their (a) counterparts in (1) are ill-formed.

1.1a) apeci-ka aphassta-nin pyenmyeng
father sick excuse
'the excuse that Father is sick'

b) *apeci-iy aphassta-nin pyenmyeng

.2a) John-i salaissta-nin cucang
is alive claim
'the claim that John is alive'

b) *John-iy salaissta-nin cucang

As expected, where the underlined NP's are relativized, (2a) and (2b) are ungrammatical:

2a) *aphassta-nin pyenmyeng-ilo John-i hakkyo-lil kyelsekhan apeci
school absent Father
*'Father who John is absent from school with the excuse that (he) is alive.'

b) *salaissta-nin cucang-lil Mary-ka mit-nin John
believe
*'John who Mary believes the claim that (he) is alive'

Ross (1967 Sec. 4.1.5) discussed a group of NP's including claim, hope, etc. which behave anomalously with respect to the CNC when they are the object NP of make or have. Thus, sentence (3) is grammatical despite its apparent violation of the CNC.

3) The money which I have hopes that the company will squander amounts to $400,000.

Such nouns as cucang 'claim', pyenmyeng 'excuse', etc. which are irregular as to the CNC are easily identifiable because those nouns also form verbs by adding -hata 'do' after them. But as is shown in (4), these NP's which can be verbalized with -hata do not exhaust the NP's which behave anomalously as to the POSS-INSERTION rule.

4a) John-i salaissta-nún somun
rumor
'the rumor that John is alive'

b) *John-ùy salaissta-nún somun

(4b) sounds better than (2b) to most speakers, probably because in (4b), the possessive-phrase is unambiguously the subject of embedded sentence. This follows from the fact that somun cannot be verbalized with -hata. (2b), however, is good if the possessive-phrase (John-ùy) is the subject of the head noun. The possibility of this interpretation is due to the fact that cucang 'claim' can be verbalized with -hata 'do'.

REFERENCES CITED

Bach, Emmon, and Robert Harms. MS. How do languages get crazy rules? Read at U.C.L.A. Conference on Historical Linguistics in the Perspective of Transformational Theory, 1969.

Kuno, Susumu. MS. Notes on Japanese grammar. Nos. 10 and 11. To appear as Report NSF-27, Mathematic linguistics and automatic translation. Cambridge, Mass.: Harvard University.

Labov, William. 1966. The social stratification of English in New York city. Washington: Center for Applied Linguistics.

_____. 1969. Contraction and variability of the English copula. Language 45/4.715-62.

_____. 1970. The study of language in its social context. To appear Stadium Generale.

Lakoff, George. 1968. Pronouns and reference. Mimeo, Harvard University.

Langacker, Ronald W. 1966. On Pronominalization and the chain of command. Modern Studies in English: Readings in Transformational Grammar, eds. by David Reibel and Sanford A. Schane. Englewood Cliffs, N.J.: Prentice Hall.

Oh, Choon-Kyu. (In preparation). Aspects of Korean syntax: Deletion in Relativization and Pronominalization. (Forthcoming Ph.D. Dissertation, University of Hawaii.)

Ross, John Robert. 1967. Constraints on variables in syntax. Cambridge, Mass.: M.I.T. Dissertation.

Thompson, Sandra Annear. 1968. Relative clauses and conjunctions. Working Papers in Linguistics, Ohio State University Research Foundation 1.80-99.

Thompson, Sandra Annear. 1970a. Relative clause structures and constraints on types of complex sentences. Working papers in Linguistics, Ohio State University Research Foundation 6.21-40.

_____. 1970b. Deep structures of relative clauses. Working Papers in Linguistics, Ohio State University Research Foundation 6.41-58.

Zwicky, Arnold M. Jr. 1968. Naturalness arguments in syntax. Papers from the Fourth Regional Meeting Chicago Linguistic Society 94-102.

20